1952

Extension of the discriminant function for evaluating high school chemistry as a prerequisite for college chemistry

Wilbur W. Sprain
Iowa State College

EXTENSION OF THE DISCRIMINANT FUNCTION

FOR EVALUATING HIGH SCHOOL CHEMISTRY AS A

PREREQUISITE FOR COLLEGE CHEMISTRY

by

Wilbur Sprain

A Dissertation Submitted to the

Graduate Faculty in Partial Fulfillment of

The Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Major Subject: Vocational Education

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

Head of Major Department

Signature was redacted for privacy.

Dean of Graduate College

Iowa State College

1952

TABLE OF CONTENTS

LIST OF TABLES

I. INTRODUCTION

The problem of predicting achievement in college courses has received much consideration during recent years. The increased number of students attending college and the difficulty encountered by these students in adapting themselves to the college environment indicate the need for investigation in this field.

The changes in high school curricula, the emphasis placed on vocational and life adjustment education rather than college preparation, and the limited number of courses offered in smaller schools result in wide differences in the scholastic preparation of students entering college. It is particularly evident, at the freshman level, that compensation should be made for the wide variation in academic background.

College counselors are faced with the problem of advising students who have difficulty in making the adjustment to college life. Until recent years counselors had little aid in determining in advance those students who might encounter trouble in making the transition from high school. Most predictions were based upon the judgment of the particular advisor rather than on any scientific data. Within the past few years, research men in education have attacked the problem on a scientific basis using appropriate statistical techniques.

Their studies have shown that probability tables can be prepared from which the chances of success may be obtained for a given student in various courses.

Emphasis in the preparation of such tables has been on the selection of students most likely to fail a particular course. The reason for attacking the problem in this manner was to provide a basis for counseling students on the advisability of choosing decelerated courses or courses not offering college credit to prepare them for work at the college level.

Many educators feel, however, that there is an equally great need for means of selecting outstanding students. If superior students could be selected prior to enrollment in a course, special sections could be offered in which material commensurate to the student's scholastic aptitude and ability would be presented.

Through early determination of studentship and celerity of academic work proportionate thereto, interest would be fostered and assimilation of subject matter would be at a maximum.

Application of the principles of prediction of scholastic achievement may be made in most subjects. Although extreme accuracy of prediction is not yet possible, considerable improvement has been made through recent developments in statistical techniques and the application of appropriate research methods.

Extension of present statistical methods was necessary to predict academic achievement in more than two categories. It is the purpose of this study to show the application of the discriminant function to predict into high, medium, and low achievement groups. Chemistry students, due to the availability of data, were used in this study to present the applicability of such an extension of discriminant analysis.

II. REVIEW OF LITERATURE

Several studies have been made which consider the factors related to successful completion of college chemistry. A study was made in 1948 by Jackson[1] at Michigan State College to determine the relationship between success in Chemistry 101, a freshman chemistry course requiring no high school chemistry, and the ability to perform the fundamental arithmetic operations and to read materials involving chemistry. All entering chemistry students were administered the ACE Psychological Examination and the Michigan State College Pre-Test.

On the basis of scores made on these tests, an attempt was made to screen students for admission to the chemistry course. Two methods of selection were tried: (1) a minimum total score on the pre-test, and (2) a minimum score for each part of the pre-test.

Neither of these methods proved satisfactory for selecting students, and Jackson[2] decided to use the discriminant function with three variables:

> (1) the raw score of Part I of the Chemistry Pre-Test consisting of twenty-five arithmetic reasoning and computation items;
> (2) the raw score on Part II of the Chemistry Pre-Test consisting of fifty-two reading items requiring no previous chemistry, but covering chemistry material;
> (3) the ACE total decile rank.

---

[1]Jackson, Robert. The Selection of Students for Freshman Chemistry by Means of Discriminant Functions. Journal of Experimental Education. 18:209-14. 1950.

[2]Ibid. p. 209.

Separate discriminant functions were derived for men and women, as Jackson believed sex difference would operate in the performance on the tests. Each sex was divided into two groups on the basis of the chemistry work, resulting in four groups which were:

(1) Males receiving grades of A, B, or C;

(2) Males receiving grades of D or F;

(3) Females receiving grades of A, B, or C;

(4) Females receiving grades of D or F.

The critical composite score for male selection was computed from the function $U = 1/2(\bar{X}_1 - \bar{X}_2)$ where $\bar{X}_1$ was the average composite score for the first group, and $\bar{X}_2$ was the average for the second group.

Any male student having a composite score less than U should receive remedial work before entering Chemistry 101, while students having a composite score equal to or greater than U would be admitted to the course.

From his results, Jackson concluded that the linear compound of the three tests distinguished significantly between students obtaining marks of A, B, or C in Chemistry 101 and those obtaining marks of D or F. He also concluded that success in Chemistry 101 was most closely related to the score on Part II of the pre-test, and least closely related to the ACE total decile rank.

Sprain[1] made a study at the Iowa State College on the prediction of first quarter mortality in general chemistry for freshman engineering and science students. For purpose of the study, survival was defined as receiving a passing mark in the regular course at the end of the quarter. The attrition group consisted of students who (1) elected a decelerated course at midterm, (2) withdrew from the course for any cause, and (3) remained in the regular course, receiving a failing mark at the end of the quarter.

The study was limited to male students in the Divisions of Engineering and Science who entered as freshmen the fall quarter of 1947. The numbers in each category were as follows:

	Credit in High School Chemistry	
	Yes	No
Engineering		
Survival	204	76
Attrition	24	90
Science		
Survival	56	17
Attrition	10	40

A discriminant function was developed for each of the four groups using as prediction variables high school average and total score on

[1]Sprain, Wilbur. Forecasting Probability of First Quarter Mortality in General Chemistry at Iowa State College. Unpublished M. S. Thesis. Ames, Iowa, Iowa State College Library. 1951.

the American Council on Education Psychological Examination. Coefficients of multiple biserial correlation were also computed to show the usefulness of the discriminant function for forecasting survival tendency. These were as follows:

Engineering students,

With high school chemistry credit $R = 0.290$;

Without high school chemistry credit $R = 0.709$;

Science students,

With high school chemistry credit $R = 0.510$;

Without high school chemistry credit $R = 0.660$.

These discriminant functions were then modified to yield sigma scores of survival tendency which could be changed into chances in 100 of survival. The discriminant functions computed from the data of 1947 students were used to predict the survival of 1950 students with reasonable accuracy.

Conclusions drawn were that entering freshmen without high school chemistry are seriously handicapped when evaluated in terms of successful completion of Chemistry 101, and that ACE-scores and high school averages are particularly useful in prediction of survival-attrition in groups without high school chemistry.

A study was made by Betts[1] on the probability of survival in the first course in chemistry for students in the Division of Agriculture at the Iowa State College. The study included 287 male students who enrolled as freshmen in the Iowa State College.

For purposes of his study, the students were divided into the following groups:

1. Students entering with credit in high school chemistry;
2. Students entering without credit in high school chemistry;
3. Students entering with credit in vocational agriculture;
4. Students entering without credit in vocational agriculture;
5. Students enrolling in Chemistry 101 with credit in high school chemistry;
6. Students enrolling in Chemistry 101 without credit in high school chemistry;
7. Students enrolling in Chemistry 101 with credit in vocational agriculture;
8. Students enrolling in Chemistry 101 without credit in vocational agriculture.

A discriminant function was derived for each of the foregoing subgroups using high school averages and ACE-scores as prediction

[1]Betts, Merle E. Probability of Mortality in First Quarter Chemistry for Students of Agriculture at Iowa State College. Unpublished M. S. Thesis. Ames, Iowa, Iowa State College Library. 1952.

variables. The first four discriminant functions were computed to predict survival-attrition at the time the student enters college. Attrition in this instance was defined as all students who drop out of college prior to enrollment in Chemistry 101, as well as those who transferred to a decelerated course or who received a failing mark in Chemistry 101.

The last four discriminant functions were computed to predict survival-attrition at the time a student enrolls in Chemistry 101. Attrition in this case was students who transferred to the decelerated course or who received a failing mark in Chemistry 101.

Betts also assumed that the tendency to survive Chemistry 101 for students entering the Division of Agriculture to be a trichotomous variable divided into (1) students who receive a passing mark in the course, (2) students who receive a failing mark or transfer to the decelerated course at midterm, and (3) those who drop out of college prior to taking Chemistry 101. Triserial correlations were computed for the various groups and ranged from 0.235 to 0.666.

Betts concluded that the total ACE score and high school average were useful in forecasting survival-attrition in Chemistry 101, regardless of the definition of attrition chosen. The usefulness of chemistry credit in high school was indicated but was less pronounced for students of agriculture, few of whom have taken Chemistry 101 during the first quarter of college, than had been found in studies of science and engineering students who had taken Chemistry 101 during the first quarter of the freshman year.

Hunter[1] undertook an investigation to determine whether having had or not having had chemistry in high school effected the final mark in Chemistry 101 at the Iowa State College. His study included male non-veteran students who were freshmen in the fall of 1947.

Using covariance analysis, with the final mark in Chemistry 101 as the criterion of achievement, and controlling on high school average and the ACE raw score, Hunter found a highly significant advantage for students who had taken high school chemistry.

Merzbacker[2] investigated the relationship between the freshman testing program and freshman chemistry marks at the San Diego State College. The coefficients of correlation between freshman chemistry final mark and ACE-Quantitative score, and between freshman chemistry final mark and the Science Section of the Iowa High School Content Examination were 0.447 and 0.464 respectively.

[1]Hunter, William A. Effect of Study of Chemistry in High School upon Achievement in Chemistry in College. Unpublished M.S. Thesis. Ames, Iowa, Iowa State College Library. 1948.

[2]Merzbacker, Claude Z. Correlation Between the Freshman Testing Program and First Semester Chemistry at San Diego State College. Journal of Chemical Education. 26:466-467. 1949.

Merzbacker also found a multiple coefficient of correlation between the Science Section of the Iowa High School Content Examination, the ACE-Quantitative score, and the freshman chemistry final mark to be 0.554. A comparison of the mean scores of the two variables revealed that the beginning chemistry students surpassed the freshman class as a whole.

The discriminant function analysis began with Fisher[1] in 1936. Since that time many studies have been made showing various applications of this discriminant function analysis. No attempt is made in this study to list references concerning the use of the discriminant function since such a review concerning its application to psychology and education was the topic for a symposium held at Harvard, and is described in the Harvard Educational Review[2].

---

[1] Fisher, R. A. The Use of Multiple Measurements in Taxonomic Problems. Annals of Eugenics. 7:179-188. 1936.

[2] Tiedman, D. V., Rulon, P. J., Bryan, J. C. The Multiple Discriminant Function - A Symposium. The Harvard Educational Review. 21:71-95. 1951.

III. METHOD OF PROCEDURE

This investigation included male students who enrolled in the Divisions of Science and Engineering at the Iowa State College during the fall quarters of 1947, 1948, 1949, and 1950. Freshmen who were not listed as entering students in the fall quarter, who were transfer students, or whose records were incomplete were eliminated from the study. The few female students were excluded.

All students included in this study registered for Chemistry 101 their first quarter at the Iowa State College. Chemistry 101 is a course in the principles of chemistry, with no prerequisite of prior chemistry. The course consists of one lecture, three recitations, and three hours of laboratory per week. Short tests are given weekly. The student's progress is evaluated at midterm on the basis of marks received on the first four examinations, and a rating of achievement in laboratory. Students making unsatisfactory progress are counseled with regard to transferring to a decelerated course, Chemistry 100A, the final decision of making such a transfer being made by the student.

The data for this study were obtained from records in the freshman chemistry office, the registrar's office, and the Iowa State College Testing Bureau. Data collected for each student included the following:

1. Final mark in Chemistry 101;
2. Credit or non-credit in high school chemistry;
3. Carnegie units of high school mathematics, for convenience multiplied by two;
4. High school grade-point average;
5. ACE-total score;
6. College division;
7. Year entered as freshman.

Stratification of the students who qualified for the study, by division, year, and whether or not they had high school chemistry is shown in Table 1.

Equal numbers or proportionality of cases are needed for the statistical treatment of this study. A table of random numbers was used to select fifty students in each of the subgroups shown in Table 1. Thus in this study only 800 of the 2299 students available were utilized.

Symbols assigned to the variables for the statistical treatment were:

X_1 = ACE total score;

X_2 = high school grade-point average;

X_3 = Carnegie units of high school mathematics.

Table 1

Division, Year, and Chemistry Background of Available Students

	Year							
	1947		1948		1949		1950	
	H.S. Chem.		H.S. Chem.		H.S. Chem.		H.S. Chem.	
Division	Yes	No	Yes	No	Yes	No	Yes	No
Engineering	282	222	292	228	158	114	193	140
Science	81	89	113	74	93	72	90	58

Statistical measures used in the treatment of the data were: the analysis of variance, the analysis of covariance, triserial r, multiple triserial R, the discriminant function, and quintiserial r. For all statistical treatment, except quintiserial r, the final marks in Chemistry 101 were divided into a trichotomy. Students who received the mark of A or B were classified in the top group, and students who received the mark of C or D were classified as the average group. Students not receiving a passing mark, those who transferred to 100A, or those who dropped the course were classified as the attrition group.

The discriminant function derived from data of the years 1947-1950 was then used on the data from the fall quarter of 1951 to test forecasting accuracy.

IV. EVALUATION WITHOUT CONTROLS

Chi square is the usual method employed to test whether the number of cases falling into various cells of a table is significantly different from what could be expected. This technique becomes rather unwieldy whenever classification has been made on more than two characteristics. The test of significance here employed has considered that the attrition-average-top tendency is a characteristic which is normally distributed. An analysis of variance was made in the usual fashion with sigma scores assigned to the attrition, to the average, and to the top groups.

In Table 2 students are classified by division, year, whether or not they had high school chemistry, and in one of the trichotomous groups of Chemistry 101 achievement.

In order to attain the sigma scores needed in the analysis, the following notations were used:

P_1 = the proportion of students in the top group;

P_2 = the proportion of students in the average group;

P_3 = the proportion of students in the attrition group;

Z_h = height of ordinate dividing the normal curve of unit area into top and average groups;

Z_c = height of ordinate dividing the normal curve of unit area into average and attrition groups;

$Z_c - Z_h$ = difference in heights of the two ordinates.

Table 2

Chemistry Achievement by Division, Year, and Prior Chemistry Credit

Group	H.S. Chem.	Division of Engineering 1947	1948	1949	1950	Sub-Total	Division of Science 1947	1948	1949	1950	Sub-Total	Total
Attrition	Yes	7	4	2	5	18	11	11	1	8	31	49
	No	28	16	14	23	81	35	21	23	27	106	187
	Both	35	20	16	28	99	46	32	24	35	137	236
Average	Yes	20	22	21	19	82	16	18	14	18	66	148
	No	18	24	31	20	93	11	17	20	16	64	157
	Both	38	46	52	39	175	27	35	34	34	130	305
Top	Yes	23	24	27	26	100	23	21	35	24	105	203
	No	4	10	5	7	26	4	12	7	7	30	56
	Both	27	34	32	33	126	27	33	42	31	133	259
Sub-Total	Yes	50	50	50	50	200	50	50	50	50	200	400
	No	50	50	50	50	200	50	50	50	50	200	400
Total		100	100	100	100	400	100	100	100	100	400	800

The mean sigma unit of attrition-average-top tendency of Chemistry 101 marks for those in the top group is $\frac{Z_h}{P_1}$, for those in the average group the mean sigma unit is $\frac{Z_c - Z_h}{P_2}$, and the mean sigma unit for those in the attrition group is $\frac{Z_c}{P_3}$.

From the data in Table 2,

$$P_1 = 0.32375;$$

$$P_2 = 0.38125;$$

$$P_3 = 0.2950.$$

The ordinates are found in a table of the normal curve and are:

$$Z_h = 0.359344;$$

$$Z_c = 0.345035;$$

$$Z_c - Z_h = 0.014309.$$

Thus, a sigma value of 1.1099 can be assigned for each member of the top group, a value of -0.03753 for each member of the average group, and a value of -1.1696 for each member of the attrition group.

It should be noted that a negative sign accompanies the sigma score whenever the mid-point of any group lies in the lower fifty percent of a distribution. Thus for the 800 students here considered, both the attrition and average groups carry a negative sign.

The sum of the standard scores (Y) for the 400 students who had high school chemistry credit may be obtained by multiplying the number of cases shown in Table 2 by the appropriate sigma score for chemistry achievement as follows:

$$(203)(1.1099) + (148)(-0.0375) + (49)(-1.1696) = 162.4528.$$

For the 400 students who had no high school chemistry the sum of the standard scores would be:

$$(56)(1.1099) + (157)(-0.03753) + (187)(-1.1696) = -162.4528.$$

The sum of squares for high school chemistry status is then obtained in the usual manner:

$$\Sigma y^2 = \frac{(162.4528)^2}{400} + \frac{(-162.4528)^2}{400} - 0 = 131.9545.$$

The correction term is zero since the sum of the standard scores is zero.

The sum of the standard scores for the 200 students in each of the years, 1947, 1948, 1949, and 1950, respectively, would be:

$$(54)(1.1099) + (65)(-0.03753) + (81)(-1.1696) = -37.2411;$$

$$(67)(1.1099) + (81)(-0.03753) + (52)(-1.1696) = 10.5064;$$

$$(74)(1.1099) + (86)(-0.03753) + (40)(-1.1696) = 32.1236;$$

$$(64)(1.1099) + (73)(-0.03753) + (63)(-1.1696) = -5.3889.$$

The sum of squares for year is then obtained in the usual manner:

$$\frac{(-37.2411)^2 + (10.5064)^2 + (32.1236)^2 + (-5.3889)^2}{200} = 12.7912.$$

The sums of squares for the other groups are found in a similar manner and are shown in Table 3. The within sums of squares for the analysis of variance is found by use of the formula indicated by Jaspen[1]:

$$N \left[\frac{Z_h^2}{P_1} + \frac{(Z_c - Z_h)^2}{P_2} + \frac{Z_c^2}{P_3} \right].$$

Substituting numerical values, the formula becomes:

$$800 \left[\frac{(0.0359344)^2}{0.32375} + \frac{(0.014309)^2}{0.38125} + \frac{(0.345035)^2}{0.2950} \right] = 642.3558.$$

The analysis of variance is shown in Table 3.

The values of "F" for the main effects of year and high school chemistry status are significant at the 1% level, and the first-order interaction of year by high school chemistry status is significant at the 5% level.

The foregoing tests of significance have been based on the assumption that the appropriate test utilizes as the error term the mean square for within subgroups. If the main effects were tested by using the interaction, whenever significant, as the appropriate error term, significant differences among years disappear, but a significant difference between those who had and those who had not had high school chemistry remains.

---

[1]Jaspen, Nathan. Serial Correlation. Psychometrika. 11:23-30. 1946.

Table 3

Analysis of Variance Without Control

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	F
Year	3	12.7912	4.2637	5.30**
High School Chemistry	1	131.9546	131.9546	164.14**
Division	1	1.5301	1.5301	1.90
Year x H.S. Chem.	3	6.6799	2.2267	2.77*
Year x Division	3	0.8820	0.2940	0.37
H.S. Chem. x Division	1	0.1933	0.1933	0.24
Year x H.S. Chem. x Division	3	1.6252	0.5417	0.67
Within	784	642.3558	0.8039	
Total	799			

From this analysis without control upon student ability, the achievement of those who had high school chemistry is higher than for those who lacked such credit from high school. Furthermore, it appears that achievement, as here measured, is not uniform from year to year. On the other hand the evidence does not suggest that a difference exists between the achievement of engineering and science students. The significant year by high school interaction implies that the usefulness of prior credit in chemistry is not uniform from year to year. More confidence could be had in the foregoing inferences if some controls of student ability were used.

V. EVALUATION WITH CONTROLS

The inferences drawn from the analysis of variance disregarded the student's aptitude to do satisfactory work in Chemistry 101. As individual differences in ability and aptitude were known to exist among the students, it was decided to control these differences by use of the analysis of covariance. For purposes of this study the following three controls were used:

1. ACE-total score (X_1);
2. High school grade-point average (X_2);
3. Carnegie units of high school mathematics (X_3).

The ACE-score was used as a scholastic aptitude control, the high school grade-point average as a prior achievement control, and the Carnegie units of high school mathematics as a mathematical ability control. The differences among groups regarding these three variables are shown in Tables 4 and 5.

An inspection of these tables reveals the general tendency of those who have had prior chemistry to excel in student ability as measured by ACE-total score, high school grade-point average, and mathematics credit, those who have had no prior chemistry. This difference may be responsible for all, or some, of the attrition-average-top tendency of those students who had high school chemistry to excel those students who did not have high school chemistry.

Table 4

Mean Values of Prediction Variables for Students With High School Chemistry Credit by Year and Division

Year		Division of Engineering			Division of Science		
		Attrition	Average	Top	Attrition	Average	Top
1947	$\bar{X}_1$	106.14	120.05	128.43	109.64	113.06	131.39
	$\bar{X}_2$	2.37	2.98	3.33	2.46	2.74	3.24
	$\bar{X}_3$	6.71	6.95	7.43	5.00	6.31	6.22
	N	7	20	23	11	16	23
1948	$\bar{X}_1$	111.00	112.18	126.58	108.82	114.05	132.81
	$\bar{X}_2$	2.00	2.55	3.26	2.23	2.37	3.03
	$\bar{X}_3$	6.25	6.59	7.16	5.82	5.55	6.57
	N	4	22	24	11	18	21
1949	$\bar{X}_1$	143.00	112.28	131.00	115.00	113.5	129.71
	$\bar{X}_2$	2.52	2.78	3.05	1.84	2.59	3.20
	$\bar{X}_3$	5.50	6.57	7.07	5.00	6.14	6.25
	N	2	21	27	1	14	35
1950	$\bar{X}_1$	108.40	115.42	131.07	108.62	110.94	127.29
	$\bar{X}_2$	2.85	2.90	3.34	2.43	2.70	3.34
	$\bar{X}_3$	5.80	7.10	6.73	5.75	6.00	7.17
	N	5	19	26	8	18	24
Total	$\bar{X}_1$	111.94	114.88	129.37	109.26	112.85	130.15
	$\bar{X}_2$	2.44	2.80	3.24	2.35	2.60	3.21
	$\bar{X}_3$	6.22	6.79	7.09	5.48	5.98	6.52
	N	18	82	100	31	66	103

Table 5

Mean Values of Prediction Variables for Students With No High School Chemistry Credit by Year and Division

Year		Division of Engineering			Division of Science		
		Attrition	Average	Top	Attrition	Average	Top
1947	$\bar{X}_1$	106.25	121.61	126.50	106.20	109.82	128.25
	$\bar{X}_2$	2.40	3.06	3.76	2.55	3.04	2.83
	$\bar{X}_3$	6.14	6.11	5.25	5.28	5.18	5.75
	N	28	18	4	35	11	4
1948	$\bar{X}_1$	102.87	114.33	124.10	93.24	115.53	121.75
	$\bar{X}_2$	2.45	2.93	3.28	2.50	2.73	3.28
	$\bar{X}_3$	6.00	6.17	6.60	5.62	6.18	6.08
	N	16	24	10	21	17	12
1949	$\bar{X}_1$	112.21	123.03	124.20	103.83	115.55	128.14
	$\bar{X}_2$	2.50	2.91	3.37	2.35	2.86	3.24
	$\bar{X}_3$	5.86	6.55	6.60	5.13	5.60	6.00
	N	14	31	5	23	20	7
1950	$\bar{X}_1$	105.56	119.15	131.29	96.00	114.00	132.57
	$\bar{X}_2$	2.59	3.10	3.15	2.76	3.09	3.30
	$\bar{X}_3$	5.61	6.30	6.57	5.22	5.63	6.14
	N	23	20	7	27	16	7
Total	$\bar{X}_1$	106.41	119.67	126.42	100.52	114.17	126.63
	$\bar{X}_2$	2.48	2.99	3.34	2.55	2.91	3.21
	$\bar{X}_3$	5.91	6.31	6.38	5.30	5.69	6.03
	N	81	93	26	106	64	30

For the analysis of covariance stratification was made on the basis of year, division, and high school chemistry status. The sample of 800 students was drawn so that there would be 400 students in each division. The 400 students in each division were subdivided into 50 students who had had high school chemistry and 50 students who had not had high school chemistry for each year of the years 1947 through 1950.

It was necessary to calculate the deviation values for all sources of variation. The raw score sums of the criterion and prediction variables, plus the sums of squares and crossproducts, are given in Table 6. These raw score values were used to compute the deviation values of the sums of squares and crossproducts of the prediction variables in the usual manner of covariance analysis and these are shown in Tables 7, 8, and 9.

The deviation form crossproducts involving the criterion, chemistry achievement, and each of the three prediction variables were calculated in a manner similar to, but not identical with that used in covariance analysis. As an example, the $\Sigma x_1 y$ for the high school chemistry main effect is shown. The equation used is:

$$\Sigma x_1 y = \frac{(\Sigma Y_h)(\Sigma X_{1h})}{N_h} + \frac{(\Sigma Y_n)(\Sigma X_{1n})}{N_n} - \frac{(\Sigma Y_T)(\Sigma X_{1T})}{N_T};$$

where the subscript h refers to having had high school chemistry, subscript n to no high school chemistry, and subscript T to the total group.

Table 6

Raw Score Sums, Sums Squares, and Crossproducts of Prediction Variables

	N	ΣX_1	ΣX_2	ΣX_3	ΣX_1^2	ΣX_2^2	ΣX_3^2	ΣX_1X_2	ΣX_1X_3	ΣX_2X_3
Engineering										
Had Chem.	200	24372	597.50	1378	3046014	1837.6232	9823	73875.72	168161	4130.77
No Chem.	200	23037	565.53	1232	2731109	1661.0395	7930	65748.49	142260	3518.38
Science										
Had Chem.	200	24241	574.48	1237	3032529	1725.8348	8027	70706.47	150964	3592.90
No Chem.	200	21761	553.06	1107	2476459	1600.0342	6403	61073.10	121197	3075.61
Total Engineering	400	47409	1163.03	2610	5777123	3508.6727	17753	139624.21	310421	7649.15
Total Science	400	46002	1127.54	2344	5508988	3325.8690	14430	131779.57	272161	6668.51
Total Had Chem.	400	48613	1171.98	2615	6078543	3563.4580	17850	144582.19	319125	7723.67
Total No Chem.	400	44798	1118.59	2339	5207568	3262.0737	14333	126821.59	263457	6593.98
Total All Students	800	93411	2290.57	4954	11286111	6824.5317	32183	271403.78	582577	14317.66

Table 7

Deviation Sums Squares of Prediction Variables Needed For Covariance Analysis

	Σy^2	Σx_1^2	Σx_2^2	Σx_3^2
Year	12.7912	3175.50	4.36552	1.7350
High School Chemistry	131.9546	18192.78	3.56312	95.2200
Division	1.5301	2474.56	1.57443	88.4450
Year x H.S. Chem.	6.6799	121.14	3.36779	9.2200
Year x Division	0.8820	531.46	0.46726	9.0950
H.S. Chem. x Division	0.1933	1638.78	0.12727	0.3200
Year x H.S. Chem. x Division	1.6252	328.39	1.62612	7.3200
Within	642.3558	352629.74	251.05154	1294.0000
Within Plus				
Year	655.1470	355805.24	255.41706	1295.7350
High School Chemistry	774.3104	370822.52	254.61466	1389.2200
Division	643.8859	355104.30	252.62597	1382.4450
Year x H.S. Chem.	649.0357	352750.88	254.41933	1303.2200
Year x Division	643.2378	353161.20	251.51880	1303.0950
H.S. Chem. x Division	642.5491	354268.52	251.17881	1294.3200
Year x H.S. Chem. x Division	643.9810	352958.13	252.67766	1301.3200

Table 8

Deviation Crossproducts Needed For Covariance Analysis
(Prediction Variables and Criterion)

	$\Sigma x_1 y$	$\Sigma x_2 y$	$\Sigma x_3 y$
Year	114.35617	-1.11405	3.50460
High School Chemistry	1549.39353	21.68339	112.09243
Division	61.53278	1.55210	11.63306
Year x H.S. Chem.	-5.71365	4.54213	5.45257
Year x Division	-2.90523	0.48786	-0.55292
H.S. Chem. x Division	17.80113	-0.16403	-0.24875
Year x H.S. Chem. x Division	11.92283	1.58147	2.32437
Within	6916.74910	209.07244	299.16396
Within Plus			
Year	7031.09527	207.95839	302.66856
High School Chemistry	8466.14263	230.75583	411.25639
Division	6978.28188	210.62454	310.79702
Year x H.S. Chem.	6911.03545	213.61457	304.61653
Year x Division	6913.84387	209.56030	298.61104
H.S. Chem. x Division	6934.55023	208.90841	298.91521
Year x H.S. Chem. x Division	6928.67193	210.65391	301.48833

Table 9

Deviation Crossproducts Needed For Covariance Analysis
(Prediction Variables)

	$\Sigma x_1 x_2$	$\Sigma x_1 x_3$	$\Sigma x_2 x_3$
Year	8.3998	-8.7075	-0.692625
High School Chemistry	254.6036	1316.1750	18.419550
Division	62.4181	467.8275	11.800425
Year x H.S. Chem.	-1.3424	13.3550	4.327850
Year x Division	-12.4473	-30.7975	0.156675
H.S. Chem. x Division	-15.0997	-22.9000	0.211000
Year x H.S. Chem. x Division	18.6299	11.3100	1.338400
Within	3633.0752	2383.12	97.744
Within Plus			
Year	3641.4750	2374.4125	97.051375
High School Chemistry	3887.6788	3699.2950	116.163550
Division	3695.4933	2850.9475	109.544425
Year x H.S. Chem.	3631.7328	2396.4750	102.071850
Year x Division	3620.6279	2352.3225	97.900675
H.S. Chem. x Division	3617.9755	2360.2200	97.955000
Year x H.S. Chem. x Division	3651.7051	2394.4300	99.0824

All values except the Y values are given in Table 6. The Y values for this equation have been previously calculated and are:

$$\Sigma Y_h = 162.4528;$$
$$\Sigma Y_n = -162.4528;$$
$$\Sigma Y_T = 0.$$

Substituting in the foregoing equation $\Sigma x_1 y$ becomes:

$$\frac{(162.4528)(48613)}{400} + \frac{(-162.4528)(44798)}{400} - 0 = 1549.3669.$$

The other needed Σxy's were obtained in the same manner.

The within deviation Σxy's were obtained as follows. The sum of the xy's for a dichotomy is equal to Nzd[1], where d is the difference

---

[1]For a dichotomy,

$$\Sigma xy = \frac{Z}{P_a}\Sigma X_a - \frac{Z}{P_b}\Sigma X_b,$$

where subscripts a and b refer to different groups, and K to the number in each group.

$$P_a = \frac{K_a}{N}, \quad P_b = \frac{K_b}{N} \qquad \therefore \quad \Sigma xy = \frac{Nz\Sigma X_a}{K_a} - \frac{Nz\Sigma X_b}{K_b};$$

$$\Sigma xy = Nz\left(\frac{\Sigma X_a}{K_a} - \frac{\Sigma X_b}{K_b}\right);$$

$$\Sigma xy = Nz(\bar{X}_a - \bar{X}_b)$$

$$\Sigma xy = Nzd$$

between two means. Jaspen[1] has shown that this Nzd for a dichotomy is a special case of a general formula. This general formula when applied to a trichotomy becomes:

$$N[z_h\bar{X}_h + (z_c - z_h)\bar{X}_c - z_c\bar{X}_l].$$

Therefore the expression within the bracket for a trichotomy corresponds to Zd for a dichotomy.

The z-values previously given in the analysis of variance section are:

$$z_h = 0.359344;$$

$$z_c = 0.345035;$$

$$z_c - z_h = 0.014309.$$

The necessary mean values for the prediction variables are shown in Table 10.

To solve for the within deviation value of $\Sigma x_1 y$, the equation using $\bar{X}_1$ values becomes:

$$N[zd] = 800[(0.359344)(129.07)-(0.014309)(115.75)-(0.345035)(104.56)]$$

$$= 6916.74910.$$

The other within deviation crossproducts involving Y were calculated in a similar manner and are shown in Table 8.

The sums of squares and sums of crossproducts for each main effect and the interactions are then separately added to the appropriate within

[1] Jaspen, op. cit.

Table 10

Mean Values of Prediction Variables
for Attrition-Average-Top Groups

Group	N	Mean ACE-Score $\bar{X}_1$	Mean H.S. Average $\bar{X}_2$	Mean Math. Units $\bar{X}_3$
Top	259	129.07	3.23	6.67
Average	305	115.75	2.83	6.24
Attrition	236	104.56	2.49	5.61
Total	800	116.76	2.86	6.19

subgroups sum of squares and crossproducts to form the "within plus" values as is done in covariance analysis, and these values are shown in Tables 7, 8, and 9.

Regression equations were then obtained by substituting "within plus" values for the main effects of year, division, high school chemistry status, and the interactions. The regression equation, in deviation form, for the three prediction variables is:

$$y = a_1x_1 + a_2x_2 + a_3x_3.$$

The necessary normal equations to solve for a_1, a_2, and a_3 are:

$$\Sigma x_1y = a_1\Sigma x_1^2 + a_2\Sigma x_1x_2 + a_3\Sigma x_1x_3;$$

$$\Sigma x_2y = a_1\Sigma x_1x_2 + a_2\Sigma x_2^2 + a_3\Sigma x_2x_3;$$

$$\Sigma x_3y = a_1\Sigma x_1x_3 + a_2\Sigma x_2x_3 + a_3\Sigma x_3^2.$$

Substitution of the within-plus-year values yields:

$$7031.09527 = 355805.24a_1 + 3641.4756a_2 + 2374.4125a_3;$$

$$207.95839 = 3641.475a_1 + 255.41706a_2 + 97.051375a_3;$$

$$302.66856 = 2374.4125a_1 + 97.051375a_2 + 1295.735a_3.$$

Simultaneous solution of the three equations produces a within-plus-year regression equation of:

$$y = 0.012834x_1 + 0.56754x_2 + 0.167559x_3.$$

The sum of squares for regression is found from the equation:

$$a_1\Sigma x_1y + a_2\Sigma x_2y + a_3\Sigma x_3y.$$

Thus, the sum of squares for the within-plus-year is:

$$(0.012834)(7031.09527) + (0.567544)(207.95839) + (0.167559)(302.66856)$$
$$= 258.9803.$$

The sum of squares of residuals for within-plus-year is found by subtracting the sum of squares for regression from the total sum of squares, Σy_1^2, which gives a sum of squares for residuals of 396.1667.

Regression equations for the other sources of variation were also calculated, from which the sums of squares for regression and their respective sums of squares for residuals were computed, and are shown in Table 11. Tests of significance for each source of variation were made exactly as in other covariance analysis, and the results are shown in Table 11.

The value of "F" for the main effects of year and high school chemistry and the first-order interaction of year by high school chemistry are significant at the 1% level. The value of "F" for the main effect of division is significant at the 5% level.

The foregoing tests of significance have been based on the assumption that the appropriate test utilizes as the error term the mean square for within subgroups. If the main effects were tested by using the interaction, whenever significant, as the appropriate error term, significant differences among years disappears, but significant differences between those who had and those who had no high school chemistry remains.

Table 11

Analysis of Covariance

Source of Variation	Degrees of Freedom	Regression Sum of Squares	Residuals Sum of Squares	Within Eliminated Sum of Squares	Within Eliminated Mean Square	F
Within Plus						
Year	3	258.9803	396.1667	11.9961	3.9987	8.13**
High School Chemistry	1	344.8513	429.4591	45.2885	45.2885	92.07**
Division	1	257.1398	386.7461	2.5755	2.5755	5.24*
Year x H.S. Chem.	3	256.7914	392.2443	8.0737	2.6912	5.57**
Year x Division	3	258.2991	384.9387	0.7681	0.2560	0.52
H.S. Chem. x Division	1	258.3214	384.2277	0.0571	0.0571	0.12
Year x H.S. Chem. x Division	3	259.7179	384.2631	0.0925	0.0308	0.06
				Mean Square		
Within Alone	781	258.1852	384.1706	0.4919		

From this analysis with control upon scholastic aptitude, prior achievement, and mathematical ability, the achievement of those who had high school chemistry is higher than for those who had no high school chemistry. The significant year main effect suggests that differences in chemistry achievement vary from year to year, even when attempts are made to control on student ability. The significant year by high school chemistry status interaction implies that the usefulness of prior credit in high school is not uniform from year to year. The difference between divisions, although significant at the 5% level, was numerically small. With as many as 800 cases available for this study the demand for significance at the 1% level did not appear too exacting.

VI. PROBABILITY TABLES FOR FORECASTING ACHIEVEMENT

The method chosen for the treatment of the data was a modification of the discriminant function technique developed by Fisher[1]. He developed the discriminant function for the purpose of finding the weights to be assigned to a series of variables that would produce maximum separation into two normally distributed groups.

This study is concerned with maximum separation, but within a single normally distributed group which has been divided into a trichotomy. The distribution has been conceived to be made up of tendencies to survive Chemistry 101.

It has been previously shown that high school grade-point averages, ACE-scores, and Carnegie units of high school mathematics vary in the attrition-average-top groups of Chemistry 101 achievement. It was not surprising that, on the average, students in the top group excelled those in the average group, who in turn excelled those in the attrition group for the prediction variables.

To establish the usefulness of these variables, when used singly to predict the attrition-average-top tendency of Chemistry 101 marks,

---

[1]Fisher, R. A. The Use of Multiple Measurements in Taxonomic Problems. Annals of Eugenics. 7:179-188. 1936.

triserial correlations were computed from Jaspen's[1] formula:

$$r_{tri} = \frac{Z_h\bar{X}_h + (Z_c - Z_h)\bar{X}_c - Z_c\bar{X}_l}{\sigma\left[\frac{Z_h^2}{P_1} + \frac{(Z_c - Z_h)^2}{P_2} + \frac{Z_c^2}{P_3}\right]} ;$$

where $\bar{X}_l$, $\bar{X}_c$, $\bar{X}_h$ = respective means of each of the segmented groups of the continuously distributed variable;

σ = standard deviation of the continuously distributed variable.

A triserial r value was calculated between ACE-scores and the attrition-average-top tendency of Chemistry 101 marks using data from the 800 students included in this study. Triserial r values were also calculated between high school averages and the attrition-average-top tendency, and between Carnegie units of high school mathematics and the attrition-average-top tendency.

Table 12 shows the total and mean scores of the ACE-score, high school grade-point average, and Carnegie units of high school mathematics for the 800 students included in this study.

Standard deviation values were calculated for each variable, and were found to be:

$\sigma_1 = 21.76845$, $\sigma_2 = 0.576783$, $\sigma_3 = 1.371749$.

Upon solution of the formula, the degree to which the trichotomous characteristic of Chemistry 101 marks are related to the variables

[1]Jaspen, op. cit.

Table 12

Mean Values Needed for Computation of Triserial r

Group	N	ΣX_1	$\bar{X}_1$	ΣX_2	$\bar{X}_2$	ΣX_3	$\bar{X}_3$
Attrition	236	24677	104.5636	588.23	2.4925	1323	5.6059
Average	305	35035	115.7541	864.88	2.8357	1903	6.2393
Top	259	33429	129.0695	837.46	3.2334	1728	6.6718

become: for ACE-score, triserial r = 0.4947; for high school grade-point average, triserial r = 0.5643; for Carnegie units of high school mathematics, triserial r = 0.3395.

These relationships are significantly different from zero when tested by the formula for t which is used for testing the significance of biserial r,

$$t = \sqrt{\frac{\left[\left(\frac{z}{\sqrt{pq}}\right)\left(r_{bis}\right)\right]^2 (N - 2)}{1 - \left[\left(\frac{z}{\sqrt{pq}}\right)\left(r_{bis}\right)\right]^2}},$$

with the exception that for triserial r the $\frac{z}{\sqrt{pq}}$ is replaced by the term within the brackets in the denominator of the formula for triserial r.

The usefulness of the combination of the three variables may be represented by a coefficient of multiple triserial correlation which is obtained by using the formula

$$R_{tri} = \frac{\sqrt{\frac{\triangle}{N}}}{\left[\frac{Z_h^2}{P_1} + \frac{(Z_c - Z_h)^2}{P_2} + \frac{Z_c^2}{P_3}\right]};$$

where $\triangle$ corresponds to the sum of squares for regression calculated from total group deviation values.

The necessary total deviation values were obtained from the original data and were as follows:

$\Sigma x_1^2 = 379092.35,$ $\Sigma x_1 x_2 = 3948.2371,$

$\Sigma x_2^2 = 266.143044,$ $\Sigma x_1 x_3 = 4129.383,$

$\Sigma x_3^2 = 1505.355,$ $\Sigma x_2 x_3 = 133.3053.$

In order to calculate the $\triangle$ corresponding to the sum of squares for regression, the following general formulas were used:

$$N(zd_1) = a_1\Sigma x_1^2 + a_2\Sigma x_1 x_2 + a_3\Sigma x_1 x_3;$$

$$N(zd_2) = a_1\Sigma x_1 x_2 + a_2\Sigma x_2^2 + a_3\Sigma x_2 x_3;$$

$$N(zd_3) = a_1\Sigma x_1 x_3 + a_2\Sigma x_2 x_3 + a_3\Sigma x_3^2.$$

Substituting,

$$(800)(8.64593637) = 379092.35a_1 + 3948.2371a_2 + 4129.383a_3;$$

$$(800)(0.26134055) = 3948.2371a_1 + 266.143044a_2 + 133.3053a_3;$$

$$(800)(0.373954956) = 4129.383a_1 + 133.3053a_2 + 1505.355a_3.$$

Solving,

$$a_1 = 0.01110713455, \qquad a_2 = 0.5614106534, \qquad a_3 = 0.1185496687.$$

The value of $\triangle$ is then,

$$800[(0.01110713455)(8.64593637) + (0.5614106534)(0.261340550)$$
$$+ (0.1185496687)(0.373954956)] = 229.66654688.$$

Therefore, $\dfrac{\triangle}{N} = \dfrac{229.6654688}{800} = 0.287083136.$

Substituting in the formula for multiple triserial R,

$$R = \frac{\sqrt{0.2870831836}}{\dfrac{(0.359344)^2}{0.32375} + \dfrac{(0.014309)^2}{0.38125} + \dfrac{(0.345035)^2}{0.2950}} = 0.6673.$$

Multiple triserial R values were also calculated for each combination of two variables and the attrition-average-top tendency of Chemistry 101 marks. An "F" test was then made to determine whether or not a significant loss was incurred by dropping a variable.

The multiple triserial R value for high school averages and ACE-scores was 0.6376, for high school averages and Carnegie units of mathematics the value was 0.6078, and for ACE-scores and Carnegie units of mathematics a value of 0.5578 was obtained. An F-value of 74.48 was obtained when ACE-scores were dropped from the three-variable regression, an F-value of 38.56 when Carnegie units of mathematics were dropped, and an F-value of 133.46 when high school averages were dropped. All three F-values are significant at the 1% level, consequently the best prediction scheme will be obtained if all three variables are used in the discriminant function.

From the analysis of covariance significant differences were noted at the 1% level for year and high school chemistry status sources of variation, and at the 5% level for the division source of variation. In the building of probability tables it was decided to ignore the significant differences among years since prediction for any group cannot include the year-to-year fluctuation. Separate tables were not developed by divisions since differences were numerically small, although significant at the 5% level. The differences were so great, however, between those who had high school chemistry and those who had not that it was necessary to develop separate equations and separate probability tables for these two groups. Such probability tables have been prepared in a trichotomy of Chemistry 101 marks, for the convenience of counselors.

Since three-way tables are extremely difficult to construct or interpret, two two-way tables were made. For the purpose of preparing each of these two-way tables a discriminant function was derived using high school averages and ACE-scores as variables, and another discriminant function was derived using high school averages and Carnegie units of mathematics as variables.

The discriminant function was used to obtain the best weights to be assigned to the variables of ACE-score and high school average for producing the maximum forecasting of the trichotomy of Chemistry 101 achievement. The discriminant function employs the solution of the following simultaneous equations:

$$Nzd_1 = a_1\Sigma x_1^2 + a_2\Sigma x_1 x_2;$$

$$Nzd_2 = a_1\Sigma x_1 x_2 + a_2\Sigma x_2^2.$$

The values used were within deviation sums of squares and cross-products for the 800 students which are shown in Tables 7 and 9. Substituting these within deviation values into the simultaneous equations they become:

$$6916.74910 = 352629.74a_1 + 3633.0752a_2;$$

$$209.07244 = 3633.0752a_1 + 251.05154a_2.$$

Upon solution the discriminant function, in deviation form, becomes:

$$v = 0.01296824020x_1 + 0.6451179225x_2.$$

The deviation form discriminant function is changed to raw score form by the addition of a constant. Thus,

$$V = 0.0129682X_1 + 0.645118X_2 + C.$$

For the purpose of preparing the probability tables two constants were computed for students who had high school chemistry and two constants were computed for students without high school chemistry. One constant was for those students not in the attrition group, and the other for those students in the top group.

The constant computed from data of the 351 students who had high school chemistry and were not in the attrition group was:

$$V - \bar{\bar{V}} = a_1(X_1 - \bar{X}_1) + a_2(X_2 - \bar{X}_2),$$

where $\bar{\bar{V}}$ is the normal deviate for the percentage of students who had high school chemistry and were not in the attrition group. The necessary mean values are given in Table 13. Upon substitution, the equation becomes:

$$V - 1.1626 = 0.0129682(X_1 - 121.533) + 0.645118(X_2 - 2.930);$$

$$V = 0.0129682X_1 + 0.645118X_2 - 2.3037.$$

Table 13

Prediction Variable Means for Students
With and Without Credit in High School Chemistry

Group		Mean ACE-Score $\bar{X}_1$	Mean H.S. Ave. $\bar{X}_2$	Mean Math. Units $\bar{X}_3$
Had H.S. Chem.	400	121.533	2.930	6.538
No H.S. Chem.	400	111.995	2.796	5.848

The foregoing equation will yield the sigma score for tendency of not being in the attrition group for any student entering the Divisions of Engineering or Science with credit in high school chemistry when his high school average and ACE-score are substituted into the discriminant function. The sigma score thus obtained is converted, by use of the normal curve table of unit area, to chances in 100 of not being in the attrition group. By subtracting the chances in 100 of not being in the attrition group from 100, the chances in 100 of being in the attrition group is found.

A constant was also computed from data of the 203 students who had high school chemistry and were in the top group. The constant obtained upon solution of the usual equation was -3.4475. The discriminant function then becomes:

$$V = 0.0129682X_1 + 0.645188X_2 - 3.4475.$$

The sigma score for tendency of being in the top group is obtained by substituting the ACE-score and high school grade-point average in the foregoing discriminant function. The sigma score thus obtained is converted by use of the normal curve table to chances in 100 of being in the top group. The chances in 100 of being in the average group are then found by subtracting from 100 the sum of the chances in 100 of being in the attrition and top group.

Table 14 was constructed with various ACE-score percentiles along the abscissa, and various high school grade-point averages along the ordinate. The chances in 100 were then computed for each ACE-score percentile division and high school grade-point average to facilitate finding the chances in 100 of attrition-average-top standing for any individual student. For instance, a student with high school chemistry who has a high school grade-point average of 3.80 and a total ACE percentile rank of 65 has 71 chances in 100 of being in the top group, 25 chances in 100 of being in the average group, and 4 chances in 100 of being in the attrition group.

The discriminant function, in raw score form, derived from the data of the 213 students who had no high school chemistry and were not in the attrition group was:

$$V = 0.0129682X_1 + 0.645118X_2 - 3.1746.$$

The discriminant function, in raw score form, derived from data of the 56 students who had no high school chemistry and were in the top group was:

$$V = 0.0129682X_1 + 0.645118X_2 - 4.3365.$$

The two foregoing equations were then solved for various ACE-scores and high school grade-point averages. The chances in 100 of being in the attrition-average-top group of Chemistry 101 marks for students without high school chemistry are shown in Table 14.

Table 14

Chances in 100 of Being in Attrition-Average-Top Group Based on High School Average and ACE-Score

		ACE Percentile Rank																			
		5		15		25		35		45		55		65		75		85		95	
		Had H.S. Chem.		Had H.S. Chem.		Had H.S. Chem.		Had H.S. Chem.		Had H.S. Chem.		Had H.S. Chem.		Had H.S. Chem.		Had H.S. Chem.		Had H.S. Chem.		Had H.S. Chem.	
H.S. Ave.	Group	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
4.00	Attrition	11	37	8	30	7	26	5	23	5	21	4	19	3	17	3	15	2	13	1	9
	Average	36	42	32	43	31	44	27	44	25	43	23	42	22	41	19	39	17	38	14	35
	Top	53	21	60	27	62	30	68	33	70	36	73	39	75	42	78	46	81	49	85	56
3.80	Attrition	14	42	10	34	8	31	7	27	6	25	5	23	4	20	4	18	3	16	2	12
	Average	38	41	35	44	35	43	30	44	29	44	27	43	25	43	22	41	20	40	16	37
	Top	48	17	55	22	57	26	63	29	65	31	68	34	71	37	74	41	77	44	82	51
3.60	Attrition	17	47	13	39	11	35	9	32	8	30	7	27	6	24	5	21	4	19	3	14
	Average	41	39	37	42	38	43	33	44	32	43	30	44	27	44	25	43	23	42	18	40
	Top	42	14	50	19	51	22	58	24	60	27	63	29	67	32	70	36	73	39	79	46
3.40	Attrition	21	52	15	44	13	40	11	37	10	34	9	31	7	28	6	25	5	22	4	17
	Average	42	37	40	40	41	42	36	42	35	43	33	44	31	44	29	44	26	44	21	42
	Top	37	11	45	16	46	18	53	21	55	23	58	25	62	28	65	31	69	34	75	41
3.20	Attrition	24	57	19	49	16	45	14	42	13	39	11	36	9	33	8	30	7	27	5	21
	Average	43	34	31	38	43	40	38	41	37	42	36	43	34	43	32	43	29	43	25	43
	Top	33	9	40	13	41	15	48	17	50	19	53	21	57	24	60	27	64	30	70	36
3.00	Attrition	29	62	22	54	19	50	17	47	15	44	14	41	12	38	10	34	9	31	6	25
	Average	43	31	43	36	45	38	41	39	40	40	38	42	36	42	35	43	32	44	28	44
	Top	28	7	35	10	36	12	42	14	45	16	48	17	52	20	55	23	59	25	66	31

Table 14
(Continued)

		ACE Percentile Rank																			
		5		15		25		35		45		55		65		75		85		95	
		Had H.S. Chem.		Had H.S. Chem.		Had H.S. Chem.		Had H.S. Chem.		Had H.S. Chem.		Had H.S. Chem.		Had H.S. Chem.		Had H.S. Chem.		Had H.S. Chem.		Had H.S. Chem.	
H.S. Ave.	Group	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
2.80	Attrition	33	67	26	59	23	55	20	52	19	49	17	46	14	43	13	39	11	36	8	29
	Average	43	27	44	33	45	35	42	37	41	38	40	40	39	40	37	42	35	43	31	44
	Top	24	6	30	8	32	10	38	11	40	13	43	14	47	17	50	19	54	21	61	27
2.60	Attrition	38	71	31	64	27	60	24	57	22	54	20	51	18	48	16	44	13	40	10	34
	Average	42	25	43	30	46	32	43	34	43	36	42	37	40	38	39	40	38	42	34	43
	Top	20	4	26	6	27	8	33	9	35	10	38	12	42	14	45	16	49	18	56	23
2.40	Attrition	43	76	35	69	32	65	29	62	26	59	24	56	21	53	19	49	16	46	12	38
	Average	40	21	43	26	45	29	43	31	44	33	43	35	42	36	41	38	40	39	37	43
	Top	17	3	22	5	23	6	28	7	30	8	33	9	37	11	40	13	44	15	51	19
2.20	Attrition	48	79	40	73	36	70	33	67	31	64	28	61	25	58	22	54	20	51	15	43
	Average	38	19	42	23	45	25	43	27	43	30	43	32	43	33	43	36	41	37	39	41
	Top	14	2	18	4	19	5	24	6	26	6	29	7	32	9	35	10	39	12	46	16
2.00	Attrition	53	83	45	78	41	74	38	71	35	69	33	66	29	63	26	59	24	56	18	49
	Average	36	15	40	19	43	22	42	25	43	26	43	28	44	30	44	33	42	34	41	38
	Top	11	2	15	3	16	4	20	4	22	5	24	6	27	7	30	8	34	10	41	13
1.80 or Below	Attrition	58	86	51	81	46	78	43	76	40	73	37	71	34	68	31	64	28	61	22	54
	Average	33	13	37	17	41	19	40	21	42	23	42	25	43	27	43	30	43	31	42	46
	Top	9	1	12	2	13	3	17	3	18	4	21	4	23	5	26	6	29	8	36	10

A student without high school chemistry credit who has a high school grade-point average of 3.80 and a total ACE percentile rank of 65 has 37 chances in 100 of being in the top group, 43 chances in 100 of being in the average group, and 20 chances in 100 of being in the attrition group. A further inspection of Table 14 reveals that the student with high school chemistry has a greater probability of being in the top group, and has less probability of being in the attrition group, than does the student without high school chemistry credit when both have the same ACE percentile rank and high school average.

The discriminant function was again used to obtain the best weights to be assigned to the variables high school grade-point average and Carnegie units of high school mathematics for producing the maximum forecasting of the trichotomy of Chemistry 101 achievement.

The needed values of within deviation sums of squares and cross-products for the 800 students are shown in Tables 7 and 9. These values were substituted into the simultaneous equations, which upon solution yielded a discriminant function of:

$$v = 0.76528x_2 + 0.17339x_3.$$

The deviation form discriminant function was changed to the raw score form by the addition of a constant. Thus,

$$V = 0.76528X_2 + 0.17339X_3 + C.$$

For the purpose of preparing a probability table, two constants were computed for students who had high school chemistry credit, and two constants were computed for students without high school chemistry credit.

The equations were:

$$V = 0.76528X_2 + 0.17339X_3 - 2.2133$$

for students who had high school chemistry credit and were not in the attrition group; and

$$V = 0.76528X_2 + 0.17339X_3 - 3.3571$$

for students who had high school chemistry credit and were in the top group.

The chances in 100 that students would be in the average group were found by subtracting from 100 the sum of the chances in 100 for the attrition and top groups.

The two foregoing equations were solved for various high school grade-point averages and Carnegie units of high school mathematics. The sigma values obtained were transmuted into chances in 100 of being in the attrition-average-top groups of Chemistry 101 marks for students who had high school chemistry credit, and are shown in Table 15. This table was constructed with various units of mathematics along the abscissa, and various high school grade-point averages along the ordinate.

Table 15

Chances in 100 of Being in Attrition-Average-Top Group Based on H.S. Average and Math. Units

H.S. Ave.		Carnegie Units of High School Mathematics																			
		1/2		1		1 1/2		2		2 1/2		3		3 1/2		4		4 1/2		5	
		H.S. Chem.		H.S. Chem.		H.S. Chem.		H.S. Chem.		H.S. Chem.		H.S. Chem.		H.S. Chem.		H.S. Chem.		H.S. Chem.		H.S. Chem.	
		Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
4.00	Attrition	15	44	13	37	9	31	6	25	4	20	3	15	2	11	1	8	1	6	--	4
	Average	40	40	35	43	32	43	29	43	24	42	20	40	16	37	13	34	9	29	7	25
	Top	45	16	52	20	59	26	65	32	72	38	77	45	82	52	86	58	90	65	93	71
3.80	Attrition	19	50	15	43	11	36	8	30	6	34	4	19	3	15	2	11	1	8	1	6
	Average	42	38	39	41	36	43	32	44	28	34	24	42	19	39	15	37	12	33	9	28
	Top	39	12	46	16	53	21	60	26	66	32	72	39	78	46	83	52	87	59	90	66
3.60	Attrition	24	56	19	49	14	42	11	35	8	39	5	23	4	18	3	14	2	11	1	8
	Average	43	34	41	38	39	41	35	43	32	34	28	44	23	42	19	40	15	36	12	32
	Top	33	10	40	13	47	17	54	22	60	27	67	33	73	40	78	46	83	53	87	60
3.40	Attrition	29	62	23	55	18	48	14	41	10	35	8	28	5	23	4	18	3	14	2	10
	Average	43	31	43	35	41	39	38	42	36	43	31	44	27	43	22	42	18	39	14	36
	Top	28	7	34	10	41	13	48	17	54	22	61	28	68	34	74	40	79	47	84	54
3.20	Attrition	34	67	28	61	22	54	18	47	13	40	10	34	7	28	5	22	4	17	2	13
	Average	43	28	43	31	43	36	40	39	39	42	35	43	31	44	27	43	22	42	18	39
	Top	23	5	29	8	35	10	42	14	48	18	55	23	62	28	68	35	74	41	80	48
3.00	Attrition	40	73	33	67	27	60	22	53	17	46	13	40	10	33	7	27	5	22	3	17
	Average	41	23	43	27	44	32	42	36	41	40	38	42	34	44	30	44	26	43	22	41
	Top	19	4	24	6	29	8	36	11	42	14	49	18	56	23	63	29	69	35	75	42

Table 15
(Continued)

H.S. Ave.		Carnegie Units of High School Mathematics: 1/2 H.S. Chem. Yes	No	1 H.S. Chem. Yes	No	1 1/2 H.S. Chem. Yes	No	2 H.S. Chem. Yes	No	2 1/2 H.S. Chem. Yes	No	3 H.S. Chem. Yes	No	3 1/2 H.S. Chem. Yes	No	4 H.S. Chem. Yes	No	4 1/2 H.S. Chem. Yes	No	5 H.S. Chem. Yes	No
2.80	Attrition	46	78	39	72	33	66	27	59	21	52	17	46	13	39	9	32	7	26	5	21
	Average	39	19	42	14	43	28	43	33	43	37	40	39	37	42	34	44	29	44	25	43
	Top	15	3	19	4	24	6	30	8	36	11	43	15	50	19	57	24	64	30	70	36
2.60	Attrition	52	82	45	77	38	71	32	65	26	59	21	52	16	45	12	38	9	32	7	26
	Average	36	16	40	20	42	25	43	29	43	33	42	37	40	40	37	42	33	43	29	44
	Top	12	2	15	3	20	4	25	6	31	8	37	11	44	15	51	20	58	25	64	30
2.40	Attrition	58	86	51	81	44	76	38	71	31	64	25	58	20	51	16	44	12	37	9	31
	Average	33	13	37	17	40	21	42	25	43	30	43	33	42	37	39	40	36	43	33	44
	Top	9	1	12	2	16	3	20	4	26	6	32	9	38	12	45	16	52	20	58	25
2.20	Attrition	64	89	57	85	50	81	43	76	37	70	30	64	25	57	20	50	15	43	11	36
	Average	29	10	34	14	38	17	41	21	42	25	44	29	43	34	41	38	39	41	37	43
	Top	7	1	9	1	12	2	16	3	21	5	26	7	32	9	39	12	46	16	52	21
2.00	Attrition	69	91	63	88	56	85	50	80	43	75	36	70	30	63	24	56	19	49	15	42
	Average	26	8	30	11	34	14	37	18	40	22	42	25	43	30	43	35	41	38	39	41
	Top	5	1	7	1	10	1	13	2	17	3	22	5	27	7	33	9	40	13	46	17
1.80 or Below	Attrition	75	94	69	91	62	88	56	84	49	80	42	74	35	68	29	62	23	55	18	48
	Average	21	6	26	8	31	11	34	14	38	18	41	23	40	37	43	31	43	35	42	39
	Top	4	--	5	1	7	1	10	2	13	2	17	3	25	5	28	7	34	10	40	13

A student with high school chemistry credit who has a high school grade-point average of 3.60, and 3 1/2 units of mathematics, has 73 chances in 100 of being in the top group, 23 chances in 100 of being in the average group, and 4 chances in 100 of being in the attrition group.

The equations were:

$$V = 0.76528X_2 + 0.17339X_3 - 3.0721$$

for students who had no high school chemistry credit and were not in the attrition group, and

$$V = 0.76528X_2 + 0.17339X_3 - 4.2340$$

for students who had no high school chemistry credit and were in the top group.

The two foregoing equations were solved for various high school grade-point averages and Carnegie units of high school mathematics. The sigma values obtained were transmuted into chances in 100 of being in the attrition-average-top groups of Chemistry 101 marks for students without high school chemistry credit and are shown in Table 15.

A student without high school chemistry credit who has a high school grade-point average of 3.60, and 3 1/2 units of mathematics, has 40 chances in 100 of being in the top group, 42 chances in 100 of being in the average group, and 18 chances in 100 of being in the attrition group.

Inspection of Table 15 reveals that the student with high school chemistry credit has a greater probability of being in the top group, and less probability of being in the attrition group, than does the student without high school chemistry credit but having the same high school grade-point average and Carnegie units of mathematics.

Some evidence of the usefulness of the discriminant function based upon information from 1947-1950 can be made by making individual predictions for each of the 1951 fall quarter engineering and science students. There were 546 such students for whom the ACE-score, high school grade-point average, Carnegie units of high school mathematics, and high school chemistry status were known.

The discriminant function was used to obtain the best weights to be assigned to the variables of ACE-score, high school grade-point average, and Carnegie units of high school mathematics, for producing the maximum forecasting of the attrition-average-top groups of Chemistry 101 achievement for each of the 546 entering freshmen in the science and engineering divisions. A constant was computed for the discriminant function for students who had high school chemistry and for students who had no high school chemistry. The discriminant function was derived from within deviation values shown in Tables 7 and 9.

The needed values were substituted into the simultaneous equations which became:

$$6916.74910 = 352629.74a_1 + 3633.0752a_2 + 2383.12a_3;$$

$$209.07244 = 3633.0752a_1 + 251.05154a_2 + 97.744a_3;$$

$$299.16396 = 2382.12a_1 + 97.744a_2 + 1294.000a_3.$$

Upon solution the discriminant function, in deviation form, became:

$$v = 0.0124395x_1 + 0.588999x_2 + 0.163793x_3.$$

The deviation form discriminant function was then converted to raw score form by the addition of a constant. A constant was computed from data of the 351 students who had high school chemistry credit and were not in the attrition group. The discriminant function in raw score form for the students with high school chemistry credit was:

$$V = 0.0124395X_1 + 0.588999X_2 + 0.163793X_3 - 3.1458.$$

A constant was also computed from data of the 213 students who had no high school chemistry credit and were not in the attrition group. The discriminant function, in raw score form, for the students without high school chemistry credit was:

$$V = 0.0124395X_1 + 0.588999X_2 + 0.163793X_3 - 3.9163.$$

The discriminant function developed from 1947 through 1950 data for students with high school chemistry was applied to each freshman of the fall quarter of 1951 with credit in high school chemistry. In a similar manner the discriminant function developed from 1947 through

1950 data for students without high school chemistry credit was applied to the 1951 fall quarter freshmen without high school chemistry credit.

Students were arranged in numerical order by their sigma scores which were found by substituting the ACE-score, high school grade-point average, and Carnegie units of high school mathematics in the appropriate discriminant function.

During the fall quarters of 1947 through 1950, 32.4 percent of the students received an A or B in Chemistry 101. A cutting percent of 32.4 was then made on the 546 students of 1951 for whom predictions were made, which represented 177 students. The marks actually received by these 177 students who had the highest sigma scores are shown in the top group of Table 16. The mark of F in Table 16 includes students who failed Chemistry 101, who dropped the course, or who transferred to Chemistry 100A.

In the fall quarters of 1947 through 1950, 38.1 percent of the students received a C or D in Chemistry 101. A cutting percent of 38.1 was then made on the 546 students of 1951 for whom predictions were made, which represented 208 students. The marks actually received by these 208 students who had the highest sigma scores after the top 177 students' sigma scores had been removed, are shown in the average group of Table 16.

Table 16

Actual Mark of Students Predicted Into an Achievement Trichotomy

Predicted Standing	Actual Mark A	B	C	D	F	Total
Top	60	71	32	9	5	177
Average	5	62	72	36	33	208
Attrition	1	8	29	23	100	161
Total	66	141	133	68	138	546

During the fall quarters of 1947 through 1950, 29.5 percent of the students had failed to make a passing mark in Chemistry 101 in the fall quarter. If a cutting percent of 29.5 is made on the 546 students for whom predictions were made, 161 students should be in the failing or attrition group. The marks actually received by these 161 students with the lowest sigma scores are shown in the attrition group of Table 16.

An inspection of Table 16 clearly reveals the usefulness of the discriminant function analysis here made for forecasting the marks which students will later make. As a measure of this relationship a coefficient of triserial correlation was computed between the sigma scores obtained from the discriminant functions, and the trichotomy of Chemistry 101 marks in the fall quarter of 1951. There were 546 students, of whom 207 received an A or B, 201 received a C or D, and 138 failed to receive a passing mark in Chemistry 101. The value of this triserial correlation coefficient was 0.7546, which is slightly higher than the multiple triserial correlation coefficient obtained with the 800 students upon which the discriminant analysis was based.

It is possible to compute a similar correlation using the marks of A, B, C, D, and F rather than the achievement trichotomy. In this case, a quintiserial coefficient of correlation is needed between the predicted sigma scores and the marks actually obtained. The formula shown by Jaspen[1] yielded a quintiserial coefficient of correlation of 0.7568.

In summary, the magnitude of the serial correlations suggest the usefulness of the discriminant analysis in the prediction of achievement in chemistry. Its usefulness is also indicated when predictions resulting therefrom are just as satisfactory when applied to a group upon whom the discriminant function is not based.

[1] Jaspen. Op. Cit.

VII. SUMMARY

It was the purpose of this study to show the application of the discriminant function to predict three categories of academic achievement. Chemistry students were used to present the applicability of such an extension of discriminant analysis.

For purposes of this study the three categories of academic achievement in Chemistry 101 were designated as the attrition group, the average group, and the top group. The attrition group consisted of students who failed to receive a passing mark in Chemistry 101, who transferred to a decelerated course, or who dropped the course. The average group consisted of students who received a mark of C or D in Chemistry 101, and the top group consisted of students who received a mark of A or B in Chemistry 101.

The study was limited to 800 of the students in the Divisions of Engineering and Science who were registered for Chemistry 101 during the fall quarters of 1947, 1948, 1949, and 1950.

Using the trichotomy of Chemistry 101 marks as the criterion of achievement, an analysis of variance indicated that the achievement of those students who had high school chemistry was higher than for those who lacked such credit in high school. From the analysis of variance it also appeared that achievement was not uniform from year to year, but no evidence was found to suggest that a difference existed between the achievement of engineering and science students.

Scholastic aptitude was controlled by the ACE-score, prior achievement by the high school grade-point average, and mathematical ability by Carnegie units of high school mathematics. An analysis of covariance indicated a highly significant difference in achievement for those students who had high school chemistry credit when compared to the students without high school chemistry credit. The analysis of covariance also indicated that achievement in chemistry varied from year to year, and that the usefulness of high school chemistry is not uniform from year to year.

Triserial correlation coefficients were computed between each variable and the attrition-average-top tendency of Chemistry 101 marks. These correlations were 0.4947, 0.5643, and 0.3395, for the ACE-score. high school grade-point average, and Carnegie units of mathematics, respectively. A multiple triserial r value was computed between the three foregoing variables and the attrition-average-top tendency of Chemistry 101 marks and was 0.6673.

A discriminant function was developed using ACE-scores and high school grade-point averages as variables, and another discriminant function was developed using high school grade-point average and Carnegie units of high school mathematics as variables.

The foregoing discriminant functions were then modified in such a way that they yielded sigma scores of attrition-average-top tendency which could be changed into chances in 100. For the convenience of counselors probability tables were prepared showing the chances in 100 of being in the attrition-average-top group of Chemistry 101 marks.

Discriminant functions based upon information from 1947-1950 and employing the ACE-score, high school grade-point average, and Carnegie units of mathematics as variables were used to make individual prediction for 546 engineering and science students who entered college the fall quarter of 1951. As a measure of the relationship between sigma scores obtained from the discriminant functions and the trichotomy of Chemistry 101 marks received in the fall quarter of 1951, a coefficient of triserial correlation was computed. The value of this triserial correlation coefficient was 0.7546, which was slightly higher than the multiple triserial correlation obtained with the 800 students upon which the discriminant analysis was based.

The magnitude of the serial correlations indicates the usefulness of the discriminant analysis in the prediction of achievement in chemistry. The usefulness of the discriminant analysis is also indicated when predictions resulting therefrom are just as satisfactory when applied to a group upon which the discriminant analysis is not based.

VIII. LITERATURE CITED

Betts, Merle E. Probability of Mortality in First Quarter Chemistry For Students of Agriculture at Iowa State College. Unpublished M.S. Thesis. Ames, Iowa, Iowa State College Library. 1952.

Fisher, R. A. The Use of Multiple Measurements in Taxonomic Problems. Annals of Eugenics. 7:179-188. 1936.

Hunter, William A. Effect of Study of Chemistry in High School Upon Achievement in Chemistry in College. Unpublished M.S. Thesis. Ames, Iowa, Iowa State College Library. 1948.

Jackson, Robert. The Selection of Students for Freshman Chemistry by Means of Discriminant Functions. Journal of Experimental Education. 18:209-214. 1950.

Jaspen, Nathan. Serial Correlation. Psychometrika. 11:23-30. 1946.

Merzbacker, Claude Z. Correlation Between the Freshman Testing Program and First Semester Chemistry at San Diego State College. Journal of Chemical Education. 26:466-467. 1949.

Sprain, Wilbur. Forecasting Probability of First Quarter Mortality in General Chemistry at Iowa State College. Unpublished M.S. Thesis. Ames, Iowa, Iowa State College Library. 1951.

Tiedman, D. V., Rulon, P. J., Bryan, J. C. The Multiple Discriminant Function - A Symposium. The Harvard Educational Review. 21:71-95. 1951.